MAKING EVALUATION DATA

Actionable

. . . from insight to impact

Ann Doucette, PhD

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

OBJECTIVES

Ann M. Doucette, PhD

Why Evaluate?

- To provide credible and useful information (evidence)
- To verify that the program/initiative is doing what was planned
- To discover challenges early to optimize outcome and impact
- To assist in prioritizing resources by identifying program components that are most effective or critical
- To reveal which functions/activities to continue and which ones to end
- To anticipate improvements needed to sustain and/or increase impact
- To assess program impact, merit, worth
- To ensure continuous quality performance – organizational improvement

The Importance of Information . . .

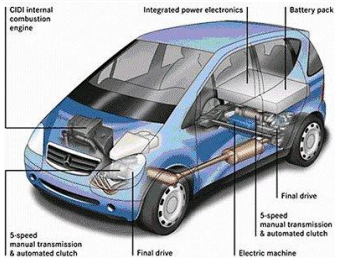
- Without Information
  - Decisions are made
    - Conjecture
    - Anecdotes – received wisdom
    - Past experience
    - Professional judgment

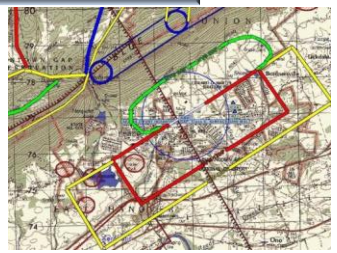
Conceptualizing Systems

Ann M. Doucette, PhD

SYSTEMS THINKING

	Easily knowable	+	Rules are given -- accepted	= Simple
Individuals and connections are equally important	+	Algorithms lead to predictable responses	+	Response is determined by the rules	= Complicated
Connections are essential; Individual representatives less critical	+	Simple rules lead to complex, adaptive, creative responses	+	Individuals have creative opportunity of responding within rules	= Complex



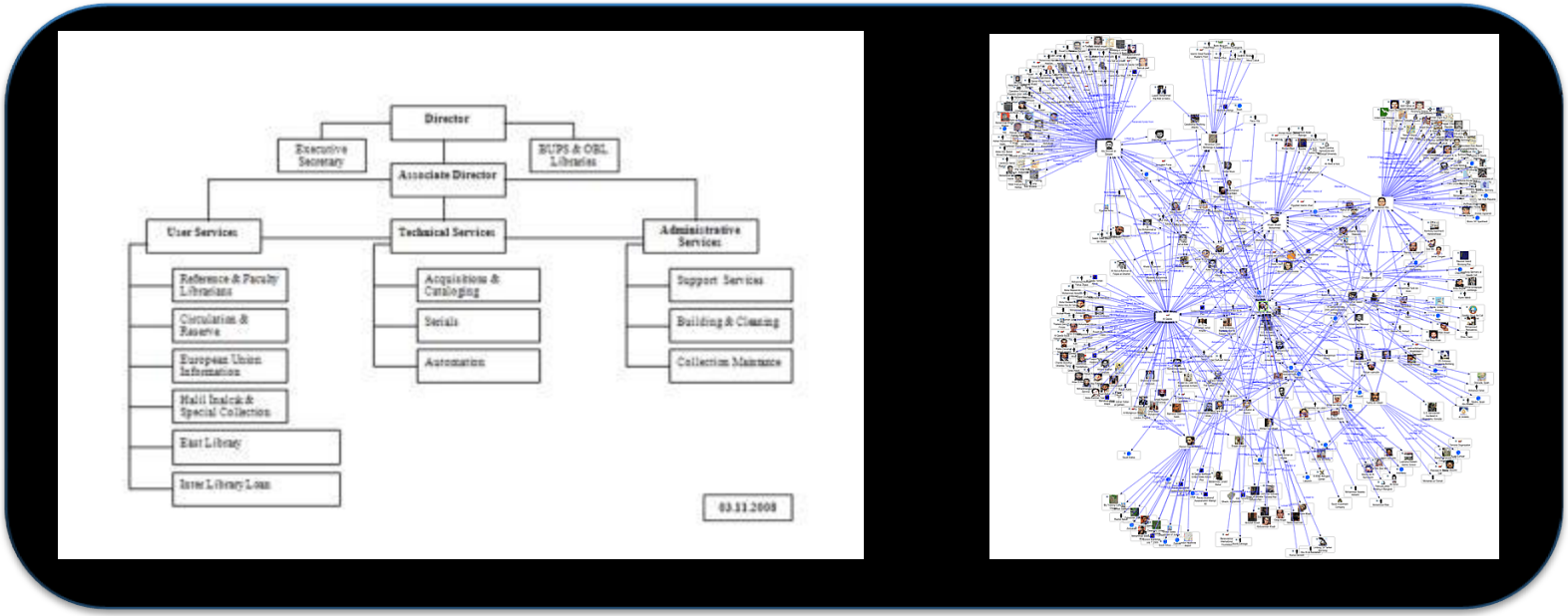


Simple **Complicated** **Complex**

At times **CHAOTIC**

Ann M. Doucette, PhD

Complexity: Human Capital, Strategy & Operations, Technology – Global Frontiers



The product of partnership networks and associated interaction.

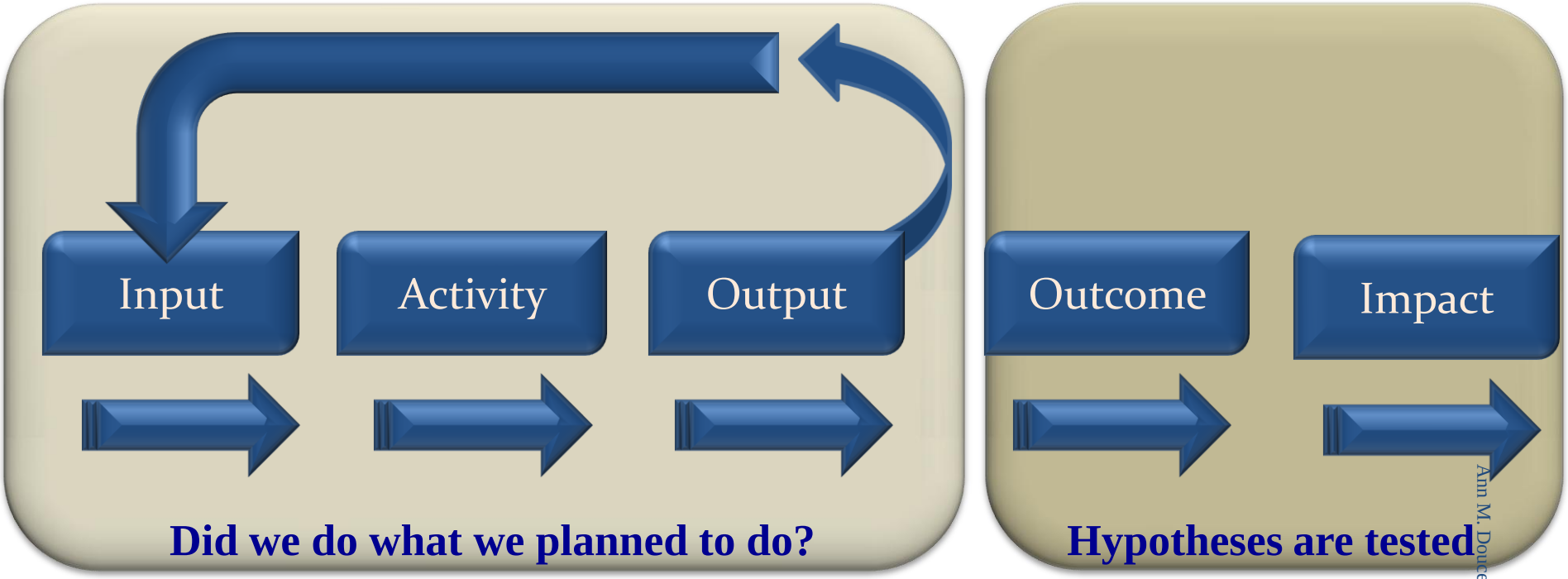
Ann M. Doucette, PhD

Systems Thinking – Evaluation Approach

- Complicated
  - Intricate number of parts
  - Lots of distinct parts
  - Evaluation perspective
    - Reductionist
      - Study of the relationship and function between/among the parts
      - Replication]
      - Predictability
- Complex
  - Intricate number of woven parts – holistic
  - The function of each part depends on the existing relationship
    - Evaluation perspective
      - Pattern description
      - Capturing dynamic evolution
      - Contextualization

Ann M. Doucette, PhD

Traditional Approach: Linear Systems



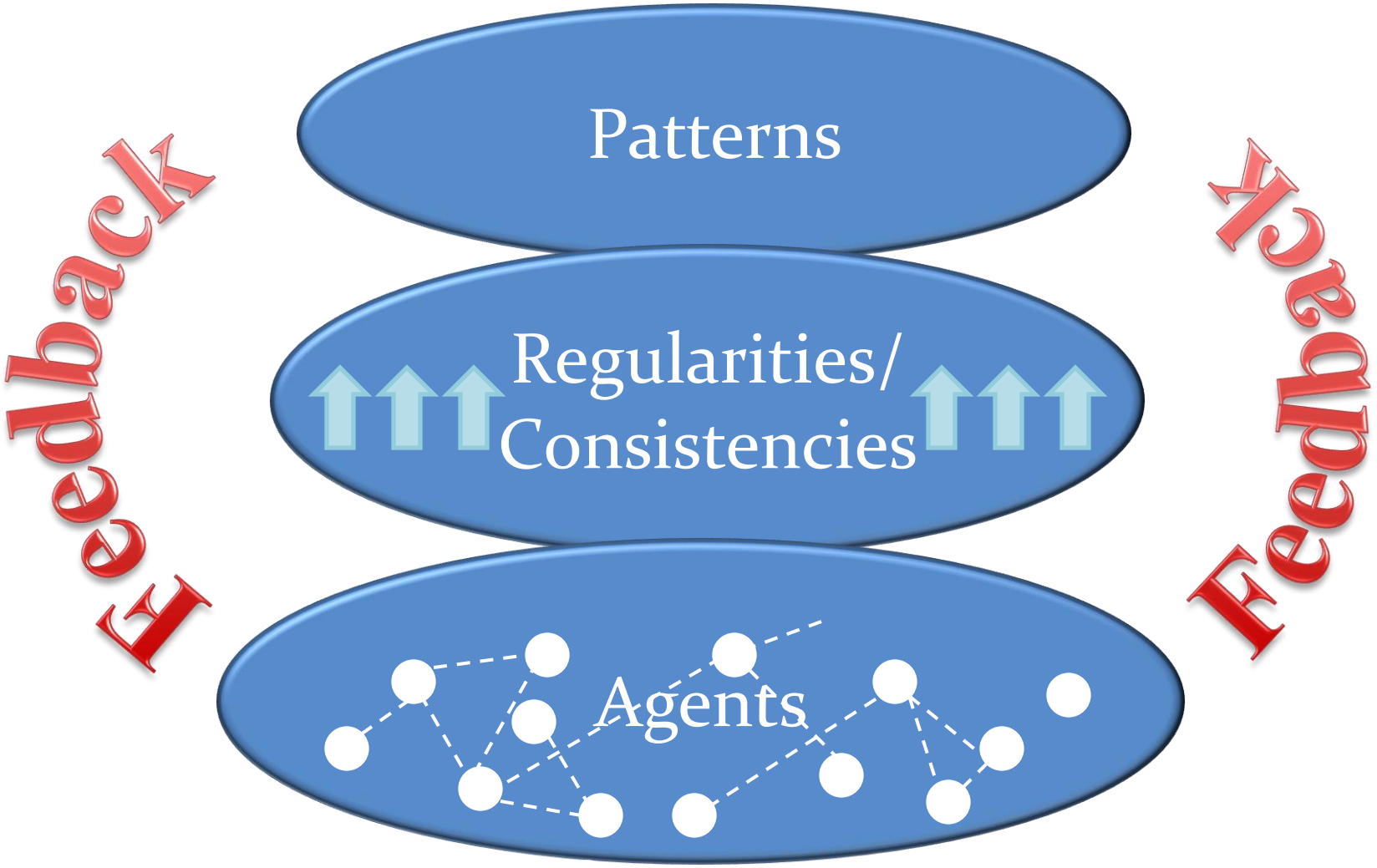
Ann M. Doucette, PhD

Complex Adaptive Systems → outputs become inputs in an iterative nonlinear manner

Traditional Approaches

- Linear – Cause → effect
- Outcomes can be understood through analysis
  - Identification of key components to explain change
  - Study of the parts
- Analysis of past events → prediction of future outcomes

Complex Adaptive Systems



Ann M. Doucette, PhD

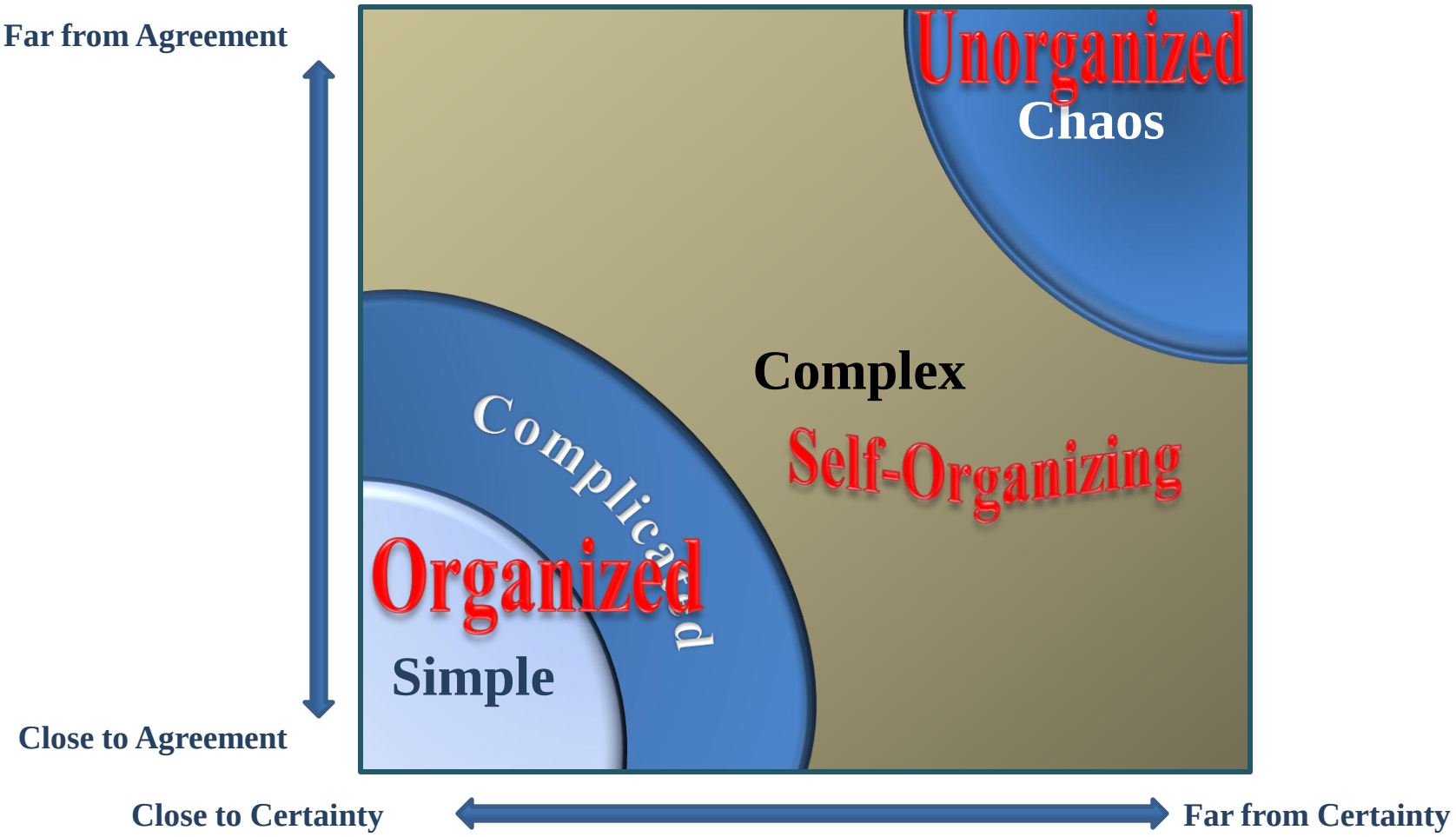
Complexity: Education

- Socio-economic status
  - Economic
  - Political
  - Business
- Gender
- Workforce development
  - Teacher training
- Job development – economic growth
- Partnership – ownership of the issues and solutions

Ann M. Doucette, PhD

Data collected as part of an interview study.

From Simple to Chaotic



Ann M. Doucette, PhD

Characteristics of Complexity

- **Emergence** – agents/representatives of the system interact in random ways
  - Interpersonal relationships
    - Within the organization
    - External to the organization – within social networks
      - Direct
      - Indirect
  - Coincidental meetings
- **Connectivity** – systems depend on interconnections and feedback
  - Dissemination of information across stakeholders is essential for the well-being of the system
  - Relationships can often be more important that the individuals

Ann M. Doucette, PhD

Characteristics of Complexity (cont.)

- **Interdependence/co-evolution** – systems operate within a specific environment (e.g., health, education, immigration, labor, etc.), however, they are also part of other systems (e.g., economic and policy environments)
  - Change in one environment precipitates the potential for change in another
  - Contradiction are potential for creative thinking
  - Small changes can have large impact (***butterfly effect*** – a ball rolling down hill may go in different ways depending on small differences in its initial position)
    - Cultural sensitivity – cultural misinterpretation
    - Exposure to rumor

Characteristics of Complexity (cont.)

- **Rules** – systems are governed by simple conventions
  - Example: Water systems (oceans, rivers, lakes, etc.) though different are governed by a simple rule – *water seeks its own level*
- **Self-organization** → consistencies and patterns
  - Example: Jazz – musical composition based on emergence (improvisation) and feedback
  - Change is often discontinuous and may appear random
  - Outcomes may be different than anticipated
- **Balance** – effectiveness versus efficiency
- **Thresholds are set** – outcomes are optimized in terms of meeting thresholds

Ann M. Doucette, PhD

CHAOS

- **Chaos**
  - Systems range from an equilibrium to chaos
  - Systems maintaining equilibrium do not respond to environmental shifts – internal or external shifts
    - Rules are imposed to control deviation
    - Attempts are made to achieve balance
  - Systems in chaos do not function as a system
    - There are no regularities, consistencies, or emerging patterns
- **Productivity and creativity are found at the edge of chaos**
  - Variety and variation are maximized
  - Potential for the emergence of new possibilities

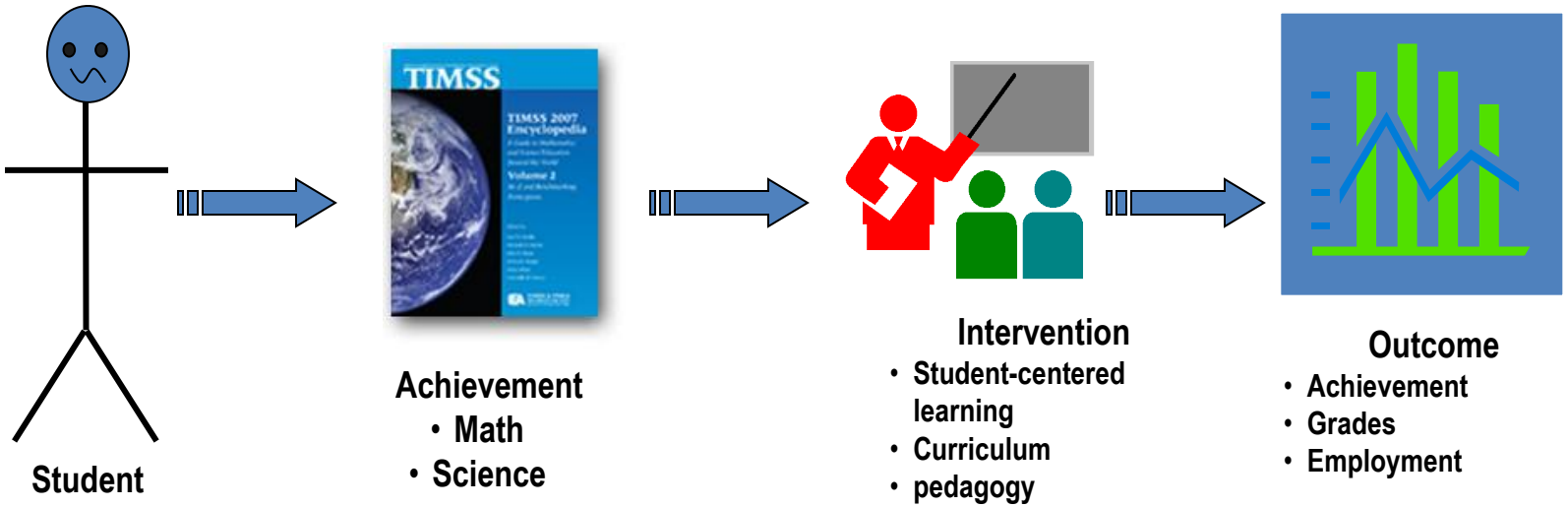
Examples: Complexity

- Health Care:
  - Medical science continues to advance
  - New emphasis on the effects patients and physician relationship, consumer decision-making, etc.
  - Rising costs, medical errors, and low quality of care continue to be problematic are thought to be symptomatic of organization problems
- Education:
  - Efforts to improve education have emphasized standardized tests in evaluating students, teachers, and schools. Standardized tests have been criticized in terms of
    - Teaching to the test
    - Unproductive uniformity among students – diminishment of creativity
- International Development:
  - To understand progress in developing nations information about economic, social, health, education, and other aspects of life is needed
  - There are inherent interdependencies which challenge development efforts
    - A focus on emerging patterns and evolutionary dynamics is needed
- Economy
  - Market value decline of corporations and the stock market
    - Mortgage policies, borrowed securities (short selling), repeal of regulations
    - Individual's ability to perceive future in terms of variable mortgages, etc.

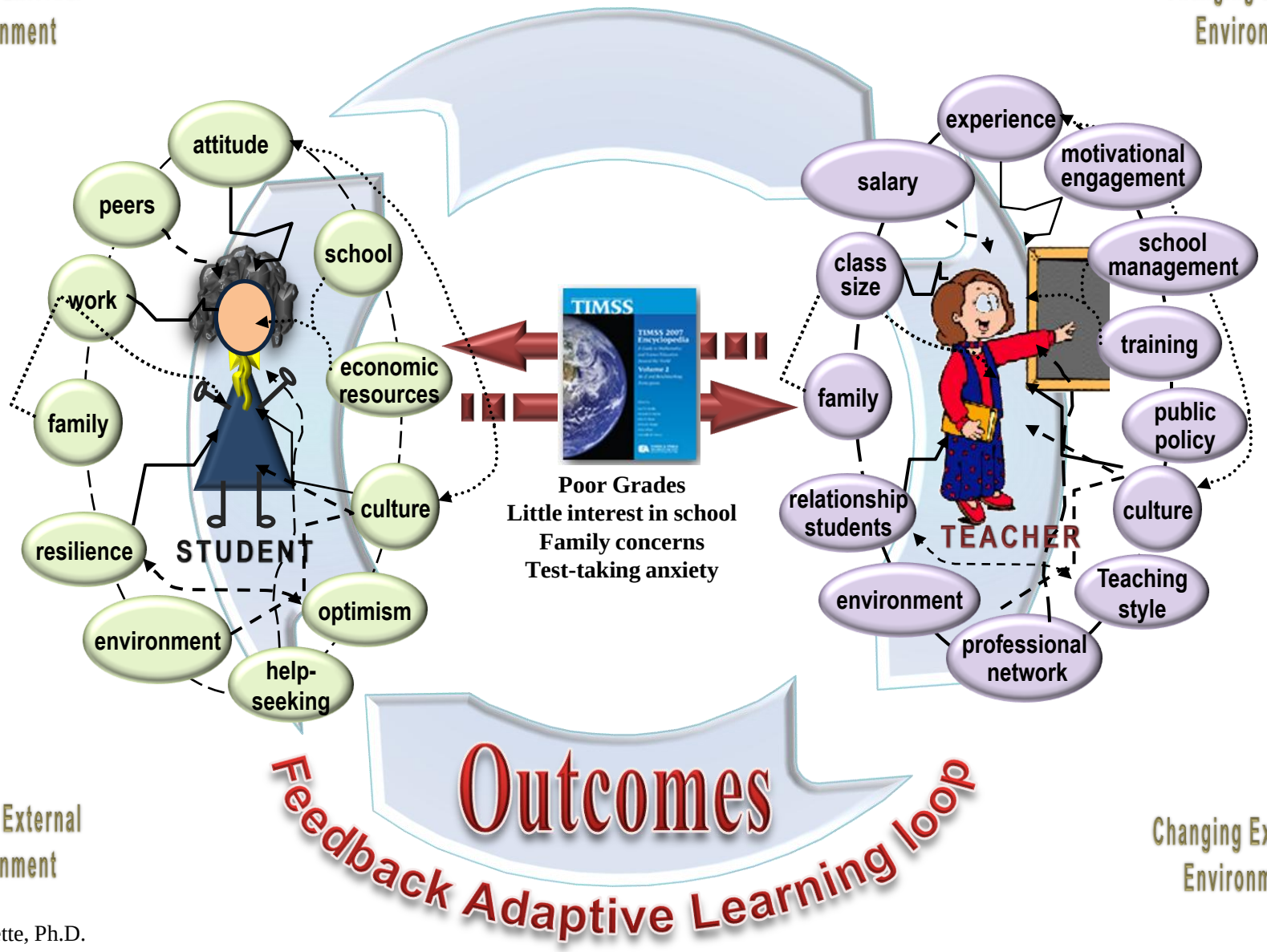
Ann M. Doucette, PhD

Characterizing the Education System

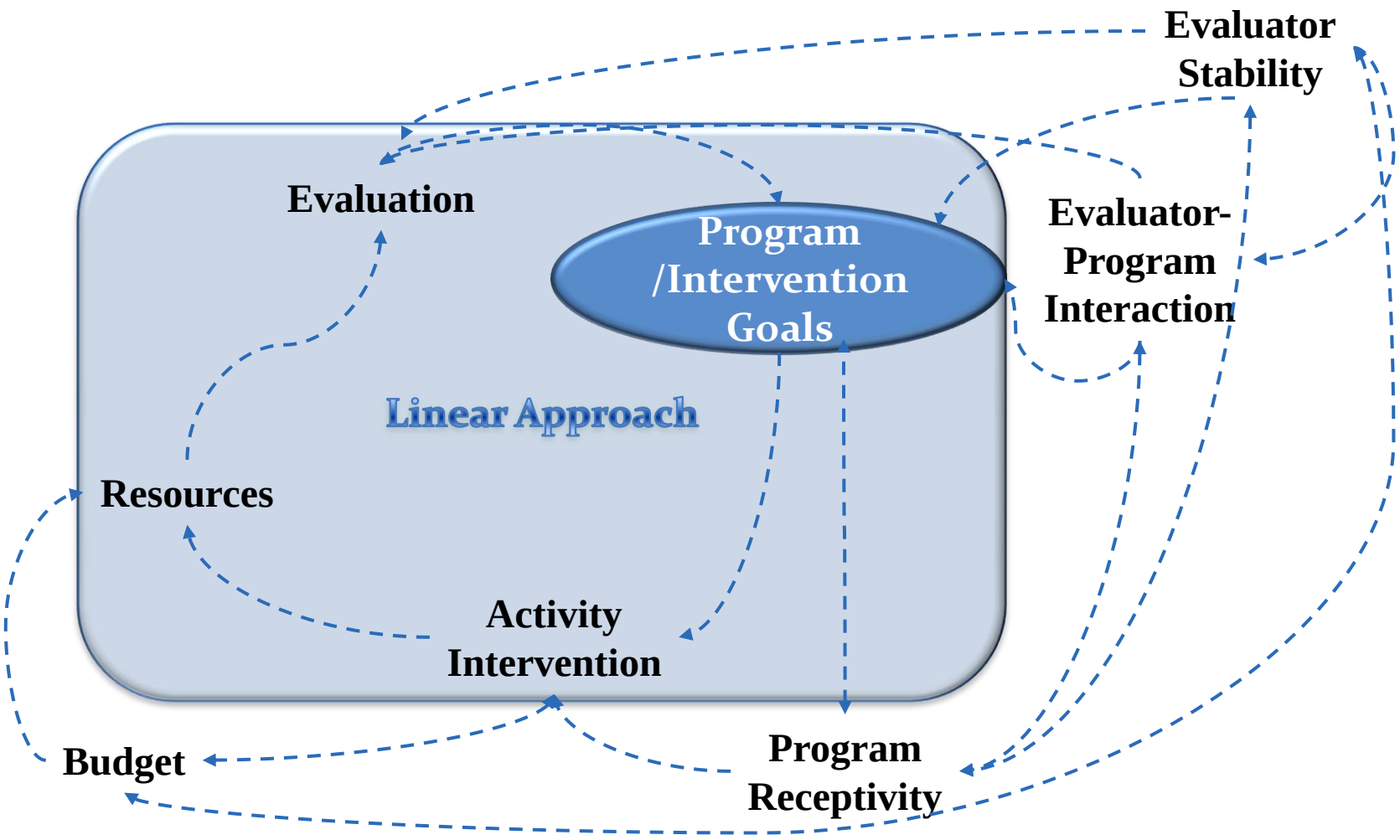
Linear – *cause and effect*



Education Reform Complexity



Evaluation: Real-world Experience



Contribution versus Attribution

- Evaluation:
  - Assessment of whether intended outcomes/impact occurred
- Contribution
  - Evaluation of how partnerships, components, etc., "contribute' to the overall outcome and impact
  - Did the program/intervention have a positive or negative affect on the outcome?
- Attribution
  - What portion of the outcome/impact can be attributed to the program/intervention?

Ann M. Doucette, PhD

Contribution Analysis

- Building a *plausible association**
  - What is the logic of the program?
  - What is the theory of change?
  - Did change take place? How is it documented?
    - What are the indicators?
  - What are alternative explanations for change?
    - What are appropriate sources of evidence?
    - What other sources of influence could be associated with change?
  - Is change stable over time?

Ann M. Doucette, PhD

*Henricks, 1996)

Theory of Change

Identifying a Theory of Change

A theory of change . . .

- Identifies a causal pathway from implementation of the intervention or program to the intended outcomes by specifying what is needed for outcomes to be achieved
- Identifies and outlines underlying assumptions which can be tested and measured
- Identified mechanisms through which change takes place
- Focuses thinking on what you are doing to what you want to achieve

Ann M. Doucette, PhD

STEPS IN BUILDING A *THEORY OF CHANGE*

1. Identifying long-term goals and the assumptions behind them
   a) Identification of connections between long term, intermediate and early outcomes
   b) Verification that the important preconditions for success have been identified
   c) Justification that the links/mechanisms between program activities and the outcomes are as expected
   d) Identification of contextual (political, economic, etc.) and environmental factors that will support or hinder progress

Steps in Building a *Theory of Change (cont.)*

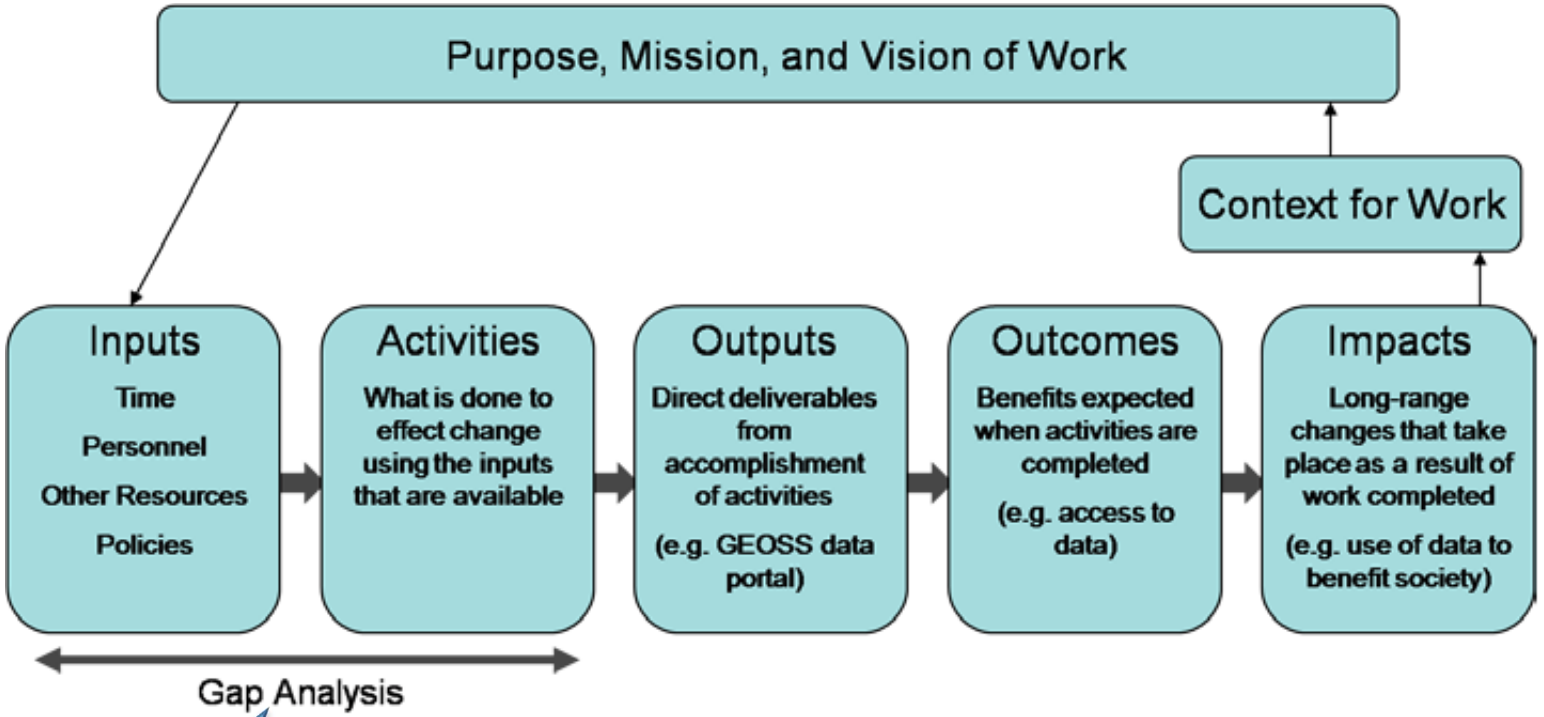
2. Backwards mapping and connect the preconditions or requirements necessary to achieve that goal
3. Identifying the interventions that your initiative will perform to create your desired change
4. Developing indicators to measure your outcomes to assess the performance of your initiative
5. Writing a narrative to explain the logic of your initiative

Using the Theory of Change Framework

- Theory of change is a framework against which the success and obstacles of your initiative will be evaluated
- Assumptions, justification, and contextual conditions believed to affect
  - Program/initiative success, information about how assumptions change, expand, or prove correct can be evaluated against the initial theory of change and modification and adaptations made throughout the program/initiative

Ann M. Doucette, PhD

Logic Model



Gap Analysis asks: Where are we? Where do we want to be?

Monitoring is an end-to-end program/project process.

Evaluation is dynamic, ongoing at all stages of the program/project.

Ann M. Doucette, PhD

Time to Expectancy
How Much Change Is Possible?

- When would change occur?
  - Will it be captured by the data collection schedule?
- Is everyone expected to make the same progress?
- What about conditions that mediate and/or moderate outcomes?
  - Is it really all about the program and the intervention?
- Does follow-up data reflect program effectiveness or other factors?

Ann M. Doucette, PhD

Approaches to Evaluation

- *Needs assessment* determines the magnitude of need and what might work to meet the need
- *Evaluability assessment* determines whether an evaluation is feasible and how stakeholders can help shape its usefulness
- *Conceptualization-focused evaluation* helps define the program or technology, the target population, and the possible outcomes
- *Implementation evaluation* monitors the fidelity of the program or technology delivery
- *Process evaluation* investigates the process of delivering the program or technology, including alternative delivery procedures

Approaches to Evaluation (cont.)

- *Developmental evaluation* focuses on what is being developed through innovative engagement
- *Outcome evaluations* investigate whether the program or technology caused demonstrable effects on specifically defined target outcomes
- *Impact evaluation* is broader and assesses the overall or net effects -- intended or unintended -- of the program or technology as a whole
- *Cost-effectiveness and cost-benefit analysis* address questions of efficiency by standardizing outcomes in terms of their dollar costs and values
- *Meta-analysis* integrates the outcome estimates from multiple studies to arrive at an overall or summary judgment on an evaluation question

Characteristics of mixed methods

- Addresses evaluation questions that focus on real-world contextual understanding
  - Cultural influences
  - Multiple perspectives
- Quantitative
  - Assess magnitude (frequency, effect size, etc.) of constructs
- Qualitative
  - Assesses the understanding and meaning of constructs
- Intentional integration of both qualitative and quantitative methods → harnessing the strengths of both
- Addressing the theoretical and philosophical frameworks in which we work

Ann M. Doucette, PhD

Advantages of Adding Qualitative Data

- Qualitative
  - Hard to measure constructs
    - Exploration of new measurement approaches
  - Complex phenomena
  - Differential interactions among measured constructs
  - Understanding of multiple perspectives
  - More precision in interpreting quantitative data

Ann M. Doucette, PhD

Mixed Method Designs

- Concurrent / Parallel
- Sequential / Exploratory
- Embedded / Nested
- Multiphase Sub-studies

Ann M. Doucette, PhD

Data and Evidence

Ann M. Doucette, PhD

Standards of Evidence

- Randomization: Assignment by chance
  - Advantages
    - Comparability of groups
    - Impartial – controlled bias
      - Reduces covariate imbalance

Standards of Evidence

- Randomization: Assignment by chance
  - Advantages
    - Comparability of groups
    - Impartial – controlled bias
      - Reduces covariate imbalance
  - **Considerations**
    - Requirement for voluntary participation may yield selection bias
      - Study conditions may affect voluntary participation
    - Protocol compliance
    - Attrition
  - **Disadvantages**
    - Highly specific populations
      - Exclusion of important, real-world participants
      - Compromised generalizability
    - Inflexible designs cannot capture **complexity** and **interconnections**
    - Smaller samples may not be generalizable to the larger population – **threat to external validity**

Ann M. Doucette, PhD

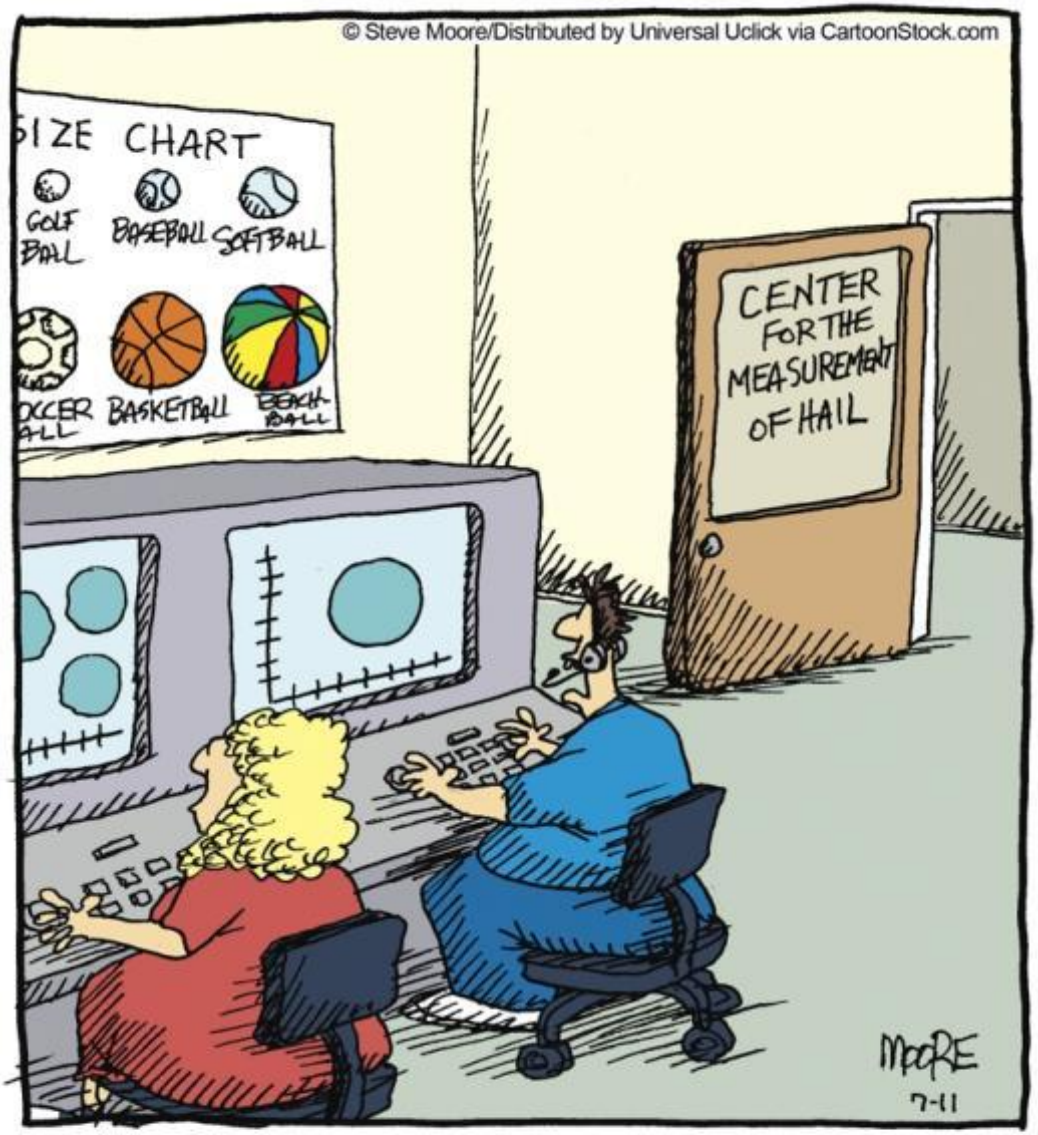
The Challenge in Identifying Equivalent Comparison Groups

- Many intervention and/or programs are non-randomly distributed
  - Participants volunteer and self-select
  - Program location is targeted
    - Participant characteristics
      - Poverty
      - Poor educational achievement
      - Unemployed
    - Geographic
- Complexity of intervention/program settings and participant characteristics make **equivalent** comparisons between "experimental" and "control" groups challenging
  - Unknown and preexisting differences → non-equivalent comparisons

Ann M. Doucette, PhD

Measurement

- Issues and Approaches



"That storm in Denver? It's confirmed. Volleyballs. What's that? ... Yes, yes. We're certain -- the size of volleyballs."

- Conjecture
- Anecdotes
- Professional Judgment?

Ann M. Doucette, PhD

What We Need . . . Data That Supports Quality Management

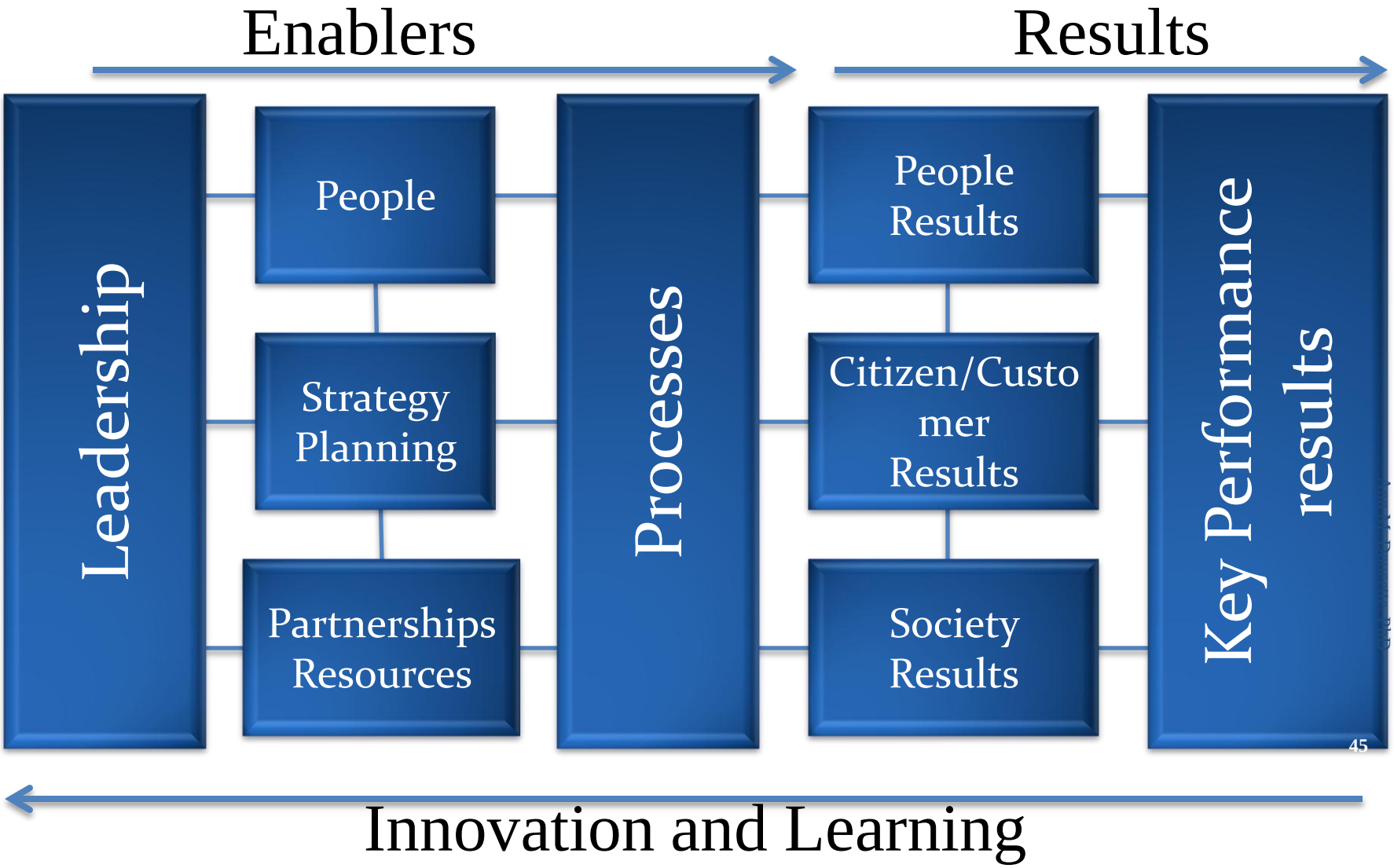
From		*To*
Compliance driven data collection	→	Outcome-based monitoring
Rule and regulation driven administration	→	Goal driven management
Best-guess decision-making	→	Data-based decision-making
Preference given to distinct professional roles	→	Cooperation across professionals is a priority
System reacts to need	→	Need is anticipated
Information is withheld	→	Information is disseminated, transparent

Ann M. Doucette, PhD

Ann Doucette, PhD, 1999

CANADIAN MANAGEMENT ACCOUNTABILITY FRAMEWORK

EUROPEAN UNION COMMON ASSESSMENT FRAMEWORK



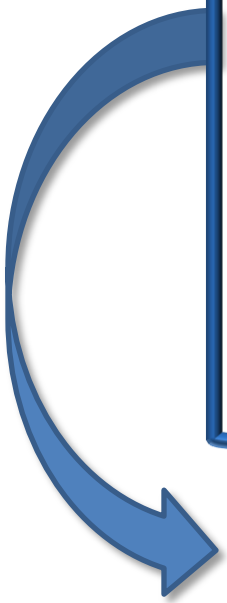
Establishing Metrics

Using the **SMART** approach. . .

- **S**pecific: explicitly targeted to the area of interest
- **M**easurable: data objectives are feasible, reliable and valid
- **A**ctionable: data findings are easy to interpret and understand, and chart a clear course for improvement over time
- **R**elevant: measure only what is important
- **T**imely: data findings are available when needed, reflecting current conditions

ANALYTIC TOOL: SWOT

- **S**trengths: attributes of the organization that are helpful to achieving the objective
- **W**eaknesses: attributes of the organization that are harmful to achieving the objective
- **O**pportunities: *external* conditions that are helpful to achieving the objective
- **T**hreats: *external* conditions that are harmful to achieving the objective



Strategy Formulation

Ann M. Doucette, PhD

Learned, E.P., Christensen, C.R., Andrews, K.E., Guth, W.D. (1965), *Business Policy: Text and Cases, Irwin, Homewood:IL.*
Weihrich, H. (1982), the TOWS matrix: a tool for situational analysis, *Long Range Planning*, 15(2), 54-66.

SWOT Examples

- **Strengths and weaknesses**
  - Resources:
    - Financial
    - Intellectual
    - Location
  - Creativity (ability to develop new approaches and new products)
  - Valuable intangible assets: intellectual capital
  - Competitive capabilities
  - Effective recruitment of talented individuals
  - Reputation
  - Motivation
  - Cultural differences
  - Bureaucracy

SWOT EXAMPLES

- **Opportunities and threats**
  - Market trends
    - Perceived need and/or importance
  - Economic conditions
  - Competition
  - Expectations of stakeholders
  - Technology
  - Public expectations
  - Criticisms by outsiders
  - Changes in markets, economy, policy
  - Policy influences (global, federal, local)
  - Regulations

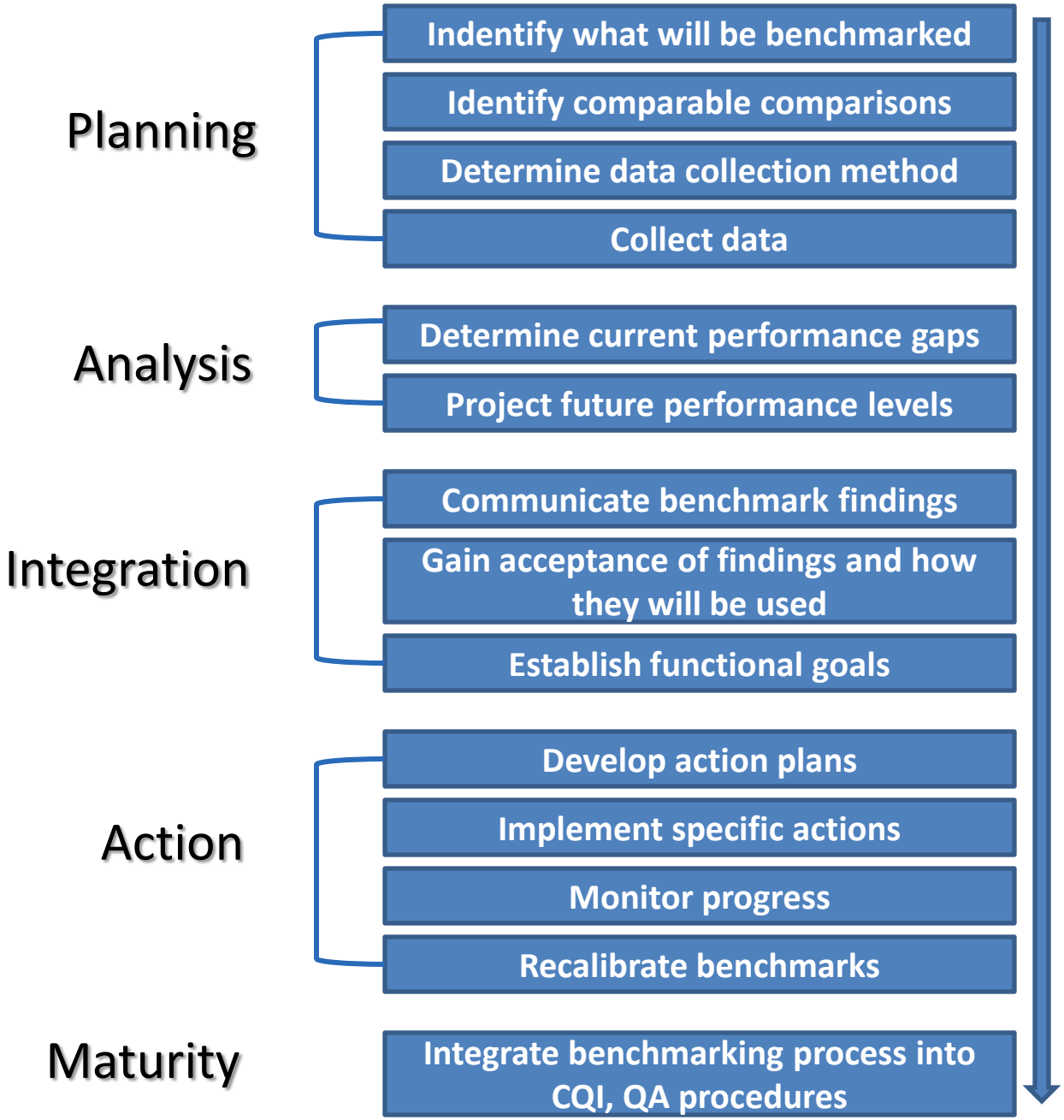
Approaches To Enhancing Quality

BENCHMARKING

- . . . Identifying the highest standard of performance, articulating the processes taken to achieve that standard, modifying and applying the processes for continued improvement

BENCHMARKING

- Strategy – what strategies are used to achieve high performance
  - Competitive
  - Targets identified to optimize performance and quality scores
    - Focus is on improvement in scores/outcomes
- Process
  - Procedures used to . . .
    - Identify need
    - Access intervention/programs
    - Exchange necessary information
    - Coordinate efforts
- Best Practices
  - Use of identified standards
  - Focus on processes and the management activities supporting them
- Cost



Adapted from: Camp, R. C., (1998). *Benchmarking: The search for industry best practices that lead to superior performance*. ASQ Press: Milwaukee, WI.

Caution In Using Benchmark Metrics

- Organization and agencies should be stable across the benchmarking interval
  - Instability may erroneously be interpreted as improvement or deterioration as a consequence of system flux
- Random variation related to sample size must be considered
  - The amount of random variation is inversely related to sample size
    - Variation is more noticeable in small samples, small agencies may be characterized as atypical as a result of sampling error
- Consider special characteristics of organizations/agencies
  - Difference from the benchmark target does not automatically indicate the need for improvement
- Consider the potential for unintended consequences
  - Staff morale
  - Team work – if benchmarking is seen as competition team work may deteriorate

Benchmarking Concerns

- Benchmarking
  - Learning from the performance of others
  - Replicating the best-practices of others
    - Following in the *foot steps* of high performers
  - No targeted focus on developing innovative solutions
- Cannot be used as a sole activity for continuous quality improvement

- *When does it pay to copy the performance of others?*

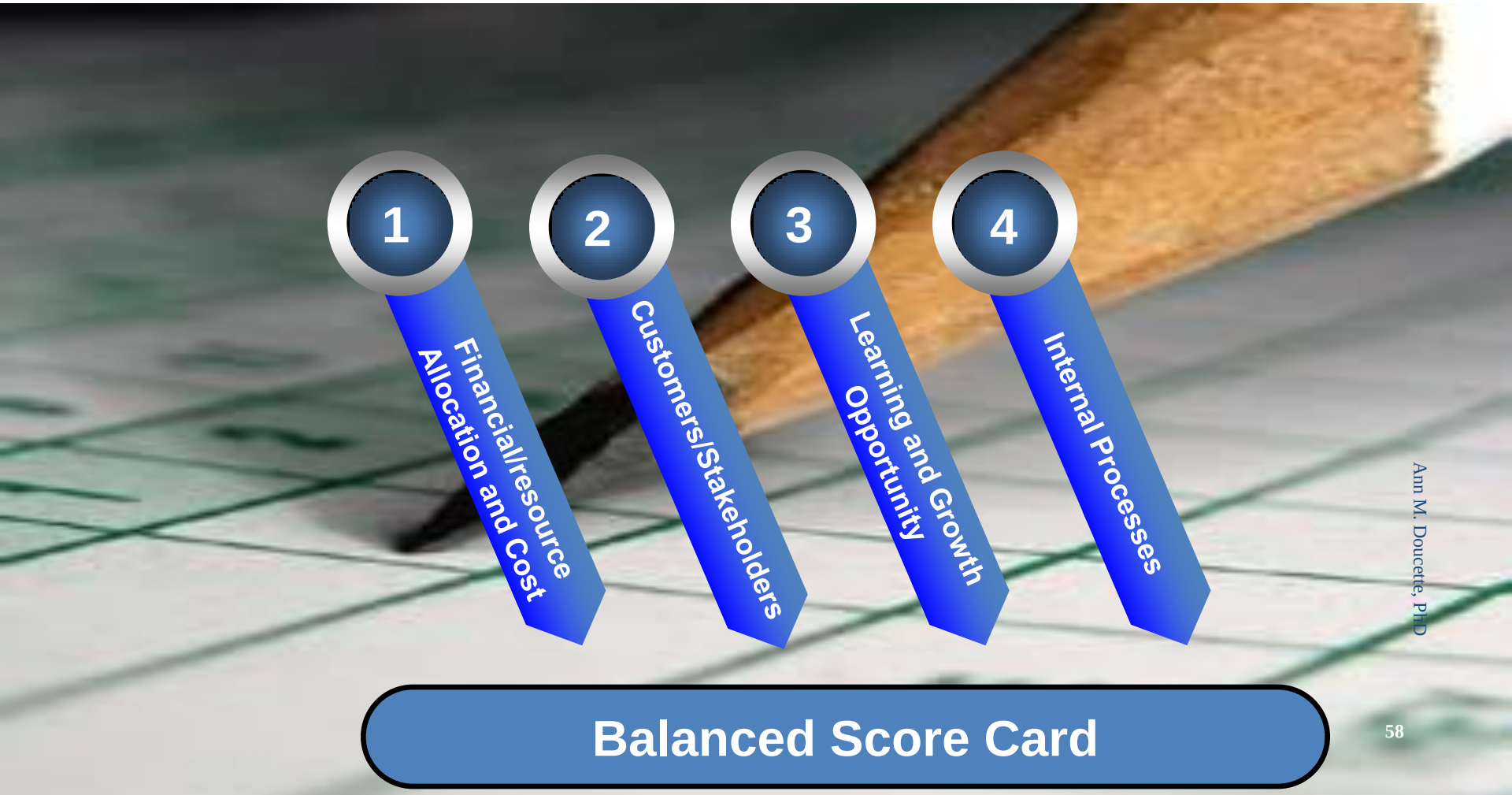
Ann M. Doucette, PhD

Making Evaluation Data *Actionable* Business Approaches

Business Frameworks and Metaphors

- **Balanced Score Card**
  - Competing interests are addressed to achieve the best overall decision to optimize outcome
  - Does improvement in one area have consequences for another?
- **Dashboard**
  - A series of indicators considered to be critical to achieving the mission
  - Focus on the end goal, with secondary monitoring of key indicators
- **Quality Standards – Quality Management**
  - Malcolm Baldrige National Quality Award
    - Approach – processes used to run an organization
    - Deployment – executing the approach/processes
    - Results – outcomes (includes process and deployment)
  - Toyota Production System – process focused
    - Value added work/activities
    - Non-value added activities
    - Waste
  - Six Sigma
    - Reduction of variability
    - Root cause and failure/effects analyses

Score Card – Addressing Complexity



Robert S. Kaplan & David P. Norton, (1996). The Balanced Scorecard : Translating Strategy

Other Performance Management Models

Capturing Complexity

Dashboard Approach

- Provides users with actionable information in an intuitive graphical format.
  - Performance data can be aggregated – rolled up into sets of key *index* performance measures (KPMs)
  - Dashboards can provide selected information that is tailored to the needs of stakeholder groups
  - KPMs can be weighted to reflect differential importance
  - Indicators can reflect actual versus projected goals

Malcolm Baldrige Performance Excellence Framework

Ann M. Doucette, PhD

Baldrige Performance Excellence: Point System

1.	**Leadership**	**120**
I.	Senior Leadership	70
II.	Governance and Societal Responsibilities	50
2.	**Strategic Planning**	**85**
I.	Strategy Development	40
II.	Strategy Implementation	45
3.	**Customer Focus**	**85**
I.	Voice of the Customer	45
II.	Customer Engagement	40

BALDRIGE PERFORMANCE EXCELLENCE: POINT SYSTEM

4.	**Measurement, Analysis, & Knowledge Management**	**90**
	I. Measurement, Analysis, and Improvement of Organizational Performance	45
	II. Management of Information, Knowledge, and Information Technology	45
5.	**Workforce Focus**	**85**
	I. Workforce Environment	40
	II. Workforce Engagement	45
6.	**Operations Focus**	**85**
	I. Work Systems	45
	II. Work Processes	40

Baldrige Performance Excellence: Point System

7.	**Results**	**450**
I.	Program and Process Outcomes	120
II.	Customer-Focused Outcomes	90
III.	Workforce-Focused Outcomes	80
IV.	Leadership and Governance Outcomes	80
V.	Financial and Market Outcomes	80

TOTAL POINTS: 1000

Ann M. Doucette, PhD

Organizational Self-Analysis

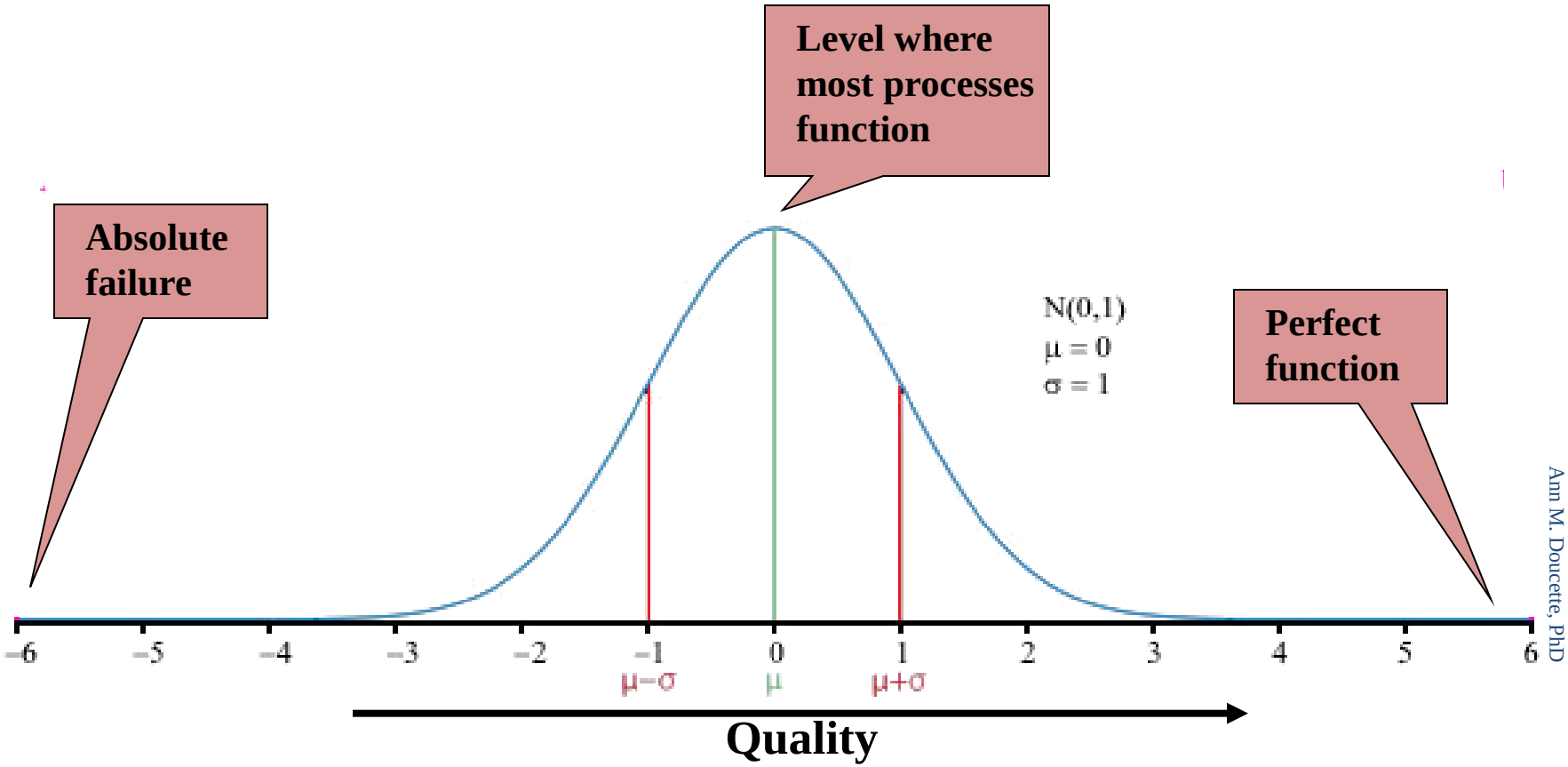
Criteria Category	Importance High, Medium, Low	For High-Importance Areas			
		Stretch (Strength) or Improvement (OFI) Goal	What Action Is Planned?	By When?	Who Is Responsible?
Category 1—Leadership					
Strength					
1.					
2.					
Opp, for Improve.					
1.					
2.					
Category 2—Strategic Planning					
Strength					
1.					
2.					
Opp, for Improve.					
1.					
2.					

- Establish goal
- Identify plan of action

Six Sigma Overview

- Six Sigma
  - Highly structured
    - Committees
    - Teams
    - Hierarchical oversight
    - Support structures
  - Resources
    - Training Six Sigma staff
    - Engagement of outside experts
  - Customer focus (e.g., patients in health care, students in school, foreign workers in training programs, etc.)
  - Error reduction → financial savings
    - The Six Sigma effort is unsuccessful if financial savings do not result from reducing error

Illustrating Six Sigma

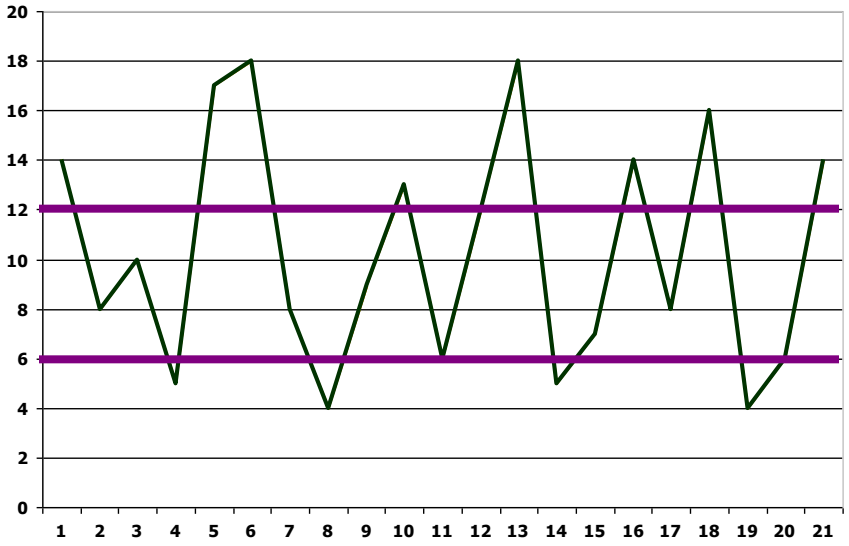


Ann M. Doucette, PhD

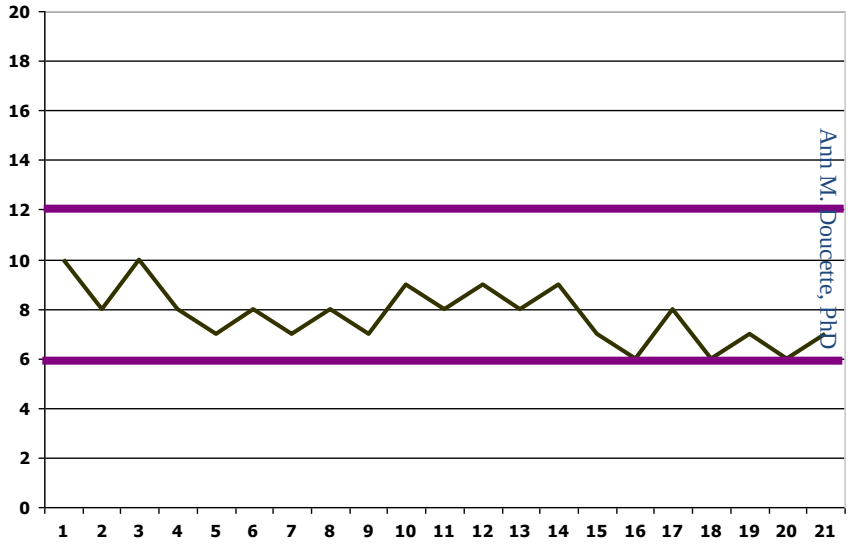
Industry standards: 6σ = 3.4 errors per 1,000,000 opportunities; 3σ sigma = 66,800 / 1,000,000; 2σ = 308,000/1,000,000

Measuring and Monitoring

- Measure
  - Identify current performance level
  - Collect data – measure performance
    - Control charts / run charts
    - Responsibility flow charts
    - Map process using sigma



Patients: pre-intervention



Patients: post-intervention

Implementing A Plan For Actionable Data

- How will data be collected?
  - Who will collect the data?
  - How often will data be collected?
- Who will use the data?
- How will the data be used?
- How will data be made available to all appropriate stakeholders?
- Which approach or method will be used?
  - How will this be decided?
- How will data be incorporated into day-to-day procedures and processes?
- Who will data be available to?
- How will administrators, managers, staff, consumers be trained to use the data?
- Will data be tied to performance incentives?
  - Who will receive the incentives?
  - How will incentives be sustained?
- What is the process for modification as goals are met?

Examples: Common Myths

- Sophisticated IT and MIS systems are needed
- Designating a quality monitor will lead to better performance
- Resources needed to increase performance and quality exist within the organization
- Demand for quality is driven externally
- IT and MIS systems should be linked to processes associated with quality
- Quality enhancement is the role of all stakeholders
- More efficient to engage within and outside resources as needed
- Anticipation of need and recognition of change can shape more effective processes

Performance Measurement Shortcoming (examples)

Performance measurement cannot . . .

- Identify all possible contributions to causality
  - Functional improvement may be attributed to the treatment model, therapeutic alliance, social connectedness, optimism/hope about the recovery process, reduction in family stress, etc.
  - If these are not measured no attribution of causality can be made.
- Assure quality of care
  - Quotas, such as time from discharge to see community-based provider says nothing about the quality of care that will be received.
- Capture the entire system
  - Measures reflect only those consumers participating/completing measures.
  - Completed data, especially follow-up data with substantial attrition cannot be generalized as representative of all consumers served.

Challenges: Quality and Performance Monitoring

The most common mistake organizations make is measuring too many variables. The next most common mistake is measuring too few.

Mark Graham Brown
Keeping Score (1996)

. . . But perhaps an even bigger mistake is to keeping doing the same thing, hoping for better results.

THANK YOU

Ann M. Doucette, PhD

Midge Smith Center for Evaluation Effectiveness
The Evaluators' Institute
The George Washington University
1922 F Street NW, Suite 407-A/B
Washington, DC 20001
Email: doucette@gwu.edu